



# A Social-aware Routing Protocol for Opportunistic Networks

Ana Cristina Kochem Vendramin, Anelise Munaretto, Myriam Regattieri Delgado, Aline Carneiro Viana, Mauro Fonseca

## ► To cite this version:

Ana Cristina Kochem Vendramin, Anelise Munaretto, Myriam Regattieri Delgado, Aline Carneiro Viana, Mauro Fonseca. A Social-aware Routing Protocol for Opportunistic Networks. Expert Systems with Applications, 2016, 34, pp.351-363. 10.1016/j.eswa.2016.01.039 . hal-01369964

**HAL Id: hal-01369964**

**<https://hal.science/hal-01369964>**

Submitted on 17 Oct 2016

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# A Social-Aware Routing Protocol For Opportunistic Networks

Ana Cristina Kochem Vendramin<sup>a,1</sup>, Anelise Munaretto<sup>a</sup>, Myriam Regattieri Delgado<sup>a</sup>, Aline Carneiro Viana<sup>b</sup>, Mauro Fonseca<sup>a</sup>

<sup>a</sup> *Computer Department (DAINF), Graduate School of Electrical Engineering and Computer Science (CPGEI), Federal University of Technology - Parana (UTFPR)*  
Av. Sete de Setembro, 3165. Rebouças. CEP 80230-901. Curitiba, Brazil.  
{criskochem,anelise,myriamdelg,maurofonseca}@utfpr.edu.br

<sup>b</sup> *INRIA Saclay - Ile de France. 4, rue Jacques Monod, 91893 Orsay Cedex France.*  
Phone: +33 (0) 1 72 92 59 98 Fax: +33 (0) 1 74 85 42 49  
aline.viana@inria.fr

---

## Abstract

This work proposes the Cultural Greedy Ant (CGrAnt) protocol to solve the problem of data delivery in opportunistic and intermittently connected networks referred to as Delay Tolerant Networks (DTNs). CGrAnt is a hybrid Swarm Intelligence-based forwarding protocol designed to address the dynamic and complex environment of DTNs. CGrAnt is based on: (1) Cultural Algorithms (CA) and Ant Colony Optimization (ACO) and (2) operational metrics that characterize the opportunistic social connectivity between wireless users. The most promising message forwarders are selected via a greedy transition rule based on local and global information captured from the DTN environment. Using simulations, we first analyze the influence of the ACO operators and CA knowledge on the CGrAnt performance. We then compare the performance of CGrAnt with the PROPHET and Epidemic protocols under varying networking parameters. The results show that CGrAnt achieves the highest delivery ratio (gains of 99.12% compared with PROPHET and 40.21% compared with Epidemic) and the lowest message replication (63.60% lower than PROPHET and 60.84% lower than Epidemic).

*Keywords:* cultural algorithms, ant colony, social analysis, forwarding protocol, intermittently connected networks.

---

<sup>1</sup>Corresponding Author: Ana Cristina Kochem Vendramin. DAINF/PPGCA/UTFPR. Av. Sete de Setembro, 3165. Rebouças. CEP 80230-901. Curitiba-PR, Brasil. Phone/Fax: +55(41)3310-4750/(41)3310-4646. criskochem@utfpr.edu.br

---

## 1. Introduction

The pervasiveness of computing devices and the emergence of new applications and cloud services are factors emphasizing the increasing need for adaptive networking solutions. In most cases, this adaptation requires the design of interdisciplinary approaches as those inspired by nature, social structures, games, and control systems. The approach presented in this paper brings together solutions from different, yet complementary domains, i.e., networking, artificial intelligence, and complex networks, and is aimed at addressing the problem of efficient data delivery in intermittently connected networks.

As mobile devices become increasingly powerful in terms of communication capabilities, the appearance of opportunistic and intermittently connected networks referred to as Delay Tolerant Networks (DTNs) is becoming a reality (Khabbaz et al., 2012; Chaintreau et al., 2007; Tournoux et al., 2011). In such networks, contacts occur opportunistically in corporate environments such as conferences sites, urban areas, or university campuses. Understanding node mobility is of fundamental importance in DTNs when designing new communication protocols that consider opportunistic encounters among nodes. In fact, it is well known in the literature that the movement of nodes in such networks is not random and is a manifestation of their routine behavior and intentions (Gonzalez et al., 2008). Together with contact-based interactions among nodes, this movement generates a mobile social network. The analysis of such mobility patterns and the understanding of how mobile nodes interact (i.e., wirelessly encounter) play a critical role in the design of solutions for DTNs.

On the other hand, given that adaptation in nature is a permanent and continuous process, we note that the dynamic and complex environment of DTNs favors the application of Swarm Intelligence (SI) methods, including approaches based on Ant Colony Optimization (ACO) (Dorigo et al., 1996) and Cultural Algorithms (CAs) (Reynolds, 1994). In fact, the environment of opportunistic DTNs presents certain features in the mobility patterns of the network nodes that can be sufficiently explored by joining CA and ACO (e.g., knowledge stored in the belief space can guide the swarms through new or already constructed paths depending on the node behavior).

Motivated by those issues, this paper proposes the use of a Cultural

Greedy Ant routing protocol, known as CGrAnt, to identify the most promising social-aware forwarders in DTNs, while profiting from SI-based paradigms. For this approach, the opportunistic and complex information (such as frequency and duration, centrality metrics, or mobility features) with respect to physical encounters between mobile nodes is gathered and favorable paths along which to forward each message are determined, while limiting data redundancy. Hence, the forwarding approach implemented by CGrAnt is adaptive and designed to match forwarding decisions to different mobility and operating conditions. Using a simulation environment, we evaluate the performance of CGrAnt under varying parameters, i.e., movement models, buffer sizes, message TTLs, simulation times, communication ranges, and transmission rates. The results confirm the satisfactory behavior of our routing protocol in key performance metrics, such as: message delivery ratio, message redundancy ratio, and message delivery delay.

The remainder of this paper is structured as follows. Section 2 provides an overview of the principles that drive our approach and its application environment. Section 3 describes the CGrAnt routing protocol in detail, and Section 4 presents the simulation environment. Sections 5 and 6 investigate how the proposed operational metrics and components affect the CGrAnt’s performance. Section 7 compares the performance of CGrAnt with two known DTN forwarding protocols under varying networking parameters, and finally, Section 8 summarizes the concluding remarks and future directions.

## 2. Rationale and Background

This section begins with an overview of the addressed problem. The main innovations and contributions are further discussed, and the state-of-the-art of forwarding in DTN environment is described, with particular attention given to approaches based on SI.

### 2.1. Problem overview

In DTNs, a fully connected multi-hop path may not exist between a sender and a receiver due to either mobility issues or varying conditions of wireless communications, thus requiring the use of specific mechanisms to ensure robustness in the data communication among nodes. The information exchange must be performed in an opportunistic fashion through so-called carry-and-forward routing techniques (Cerf et al., 2007). The nodes may

need to store messages from other nodes in their buffers for long periods of time and carry these messages until a forwarding opportunity arises (Cerf et al., 2007). Additionally, message replication may be necessary to increase the probability of successfully delivered messages. However, certain problems exist in a limited resource scenario: replications are undesirable because they compete with valid data messages in the paths toward a destination, and the storage of neighbors' data messages can be a problem due to limited buffer sizes.

The problem of routing in DTNs can thus, be modeled as a multimodal optimization problem attempting to find not just one solution but a set of solutions (i.e., multiple paths between two nodes). The finite set of possible solutions (i.e., paths formed by a sequence of nodes in which each node permutation generates a new solution) characterizes the routing in DTNs as a combinatorial problem. The problem can be also modeled as a dynamic state because the search space characteristics and the location and value of the solutions will change over time. The problem of routing in DTNs presents, therefore, a complex challenge, with several aspects still unexplored by most approaches described in the literature. Therefore, an updated consideration of the DTN dynamics is necessary and can be accomplished by periodically analyzing the neighbor information and selecting more than one path along which to forward each message while limiting message redundancy. The dynamicity and complex premises of DTNs characterize it as an environment favorable for the application of SI algorithms, including ACO and CA (Dorigo et al., 1996; Reynolds, 1994).

## 2.2. CGrAnt in a Nutshell

In view of the problem discussed in the previous session, we propose the CGrAnt protocol as a solution to the problem of identifying a set of good nodes along which to route each message in DTNs. To increase the reliability in such dynamic networks, choosing the best path for routing of messages should not be the main goal of a routing protocol. Indeed, it is equally important to maintain a diversity of paths and avoid convergence to only one or a few solutions.

CGrAnt can be defined as a hybrid SI system based on (1) CA and ACO and (2) operational metrics that characterize the opportunistic social connectivity between nodes. To adapt to the large topology variations encountered by a DTN and to reduce latency in message delivery, the following modifications are incorporated into CGrAnt that differentiate it from traditional

108 SI-based protocols. (1) The SI control messages named Forward Ants (FAs),  
 109 responsible for the path construction, are encapsulated into the data mes-  
 110 sages. (2) The number of FAs created and forwarded is dynamically defined  
 111 according to the knowledge stored in the nodes. (3) CGrAnt adopts a greedy  
 112 ACO transition rule that is similar to the deterministic transition rule pro-  
 113 posed in the Ant Colony System (ACS) (Dorigo and Gambardella, 1997).  
 114 Nevertheless, instead of using both probabilistic and deterministic rules as  
 115 in ACS, CGrAnt uses only the greedy transition rule, which considers the  
 116 heuristic function and/or pheromone concentration, to forward messages to  
 117 the most promising node(s), and/or to exploit previously found good solu-  
 118 tions. The search space exploration is still provided by the DTNs dynamics.  
 119 (4) Instead of using time-based pheromone evaporation, CGrAnt performs  
 120 an event-driven evaporation, which only occurs if a node detects that a new  
 121 path toward a destination has been found. Thus, allowing redundant paths  
 122 becomes more important than converging to the best path. (5) Because  
 123 there is no central element in DTNs, the knowledge stored in the CA belief  
 124 space is distributed among network nodes. (6) The information exchanged  
 125 between the belief and population spaces always occurs in a distributed man-  
 126 ner intermediated by the CGrAnt operational metrics. The ACO and CA  
 127 modifications seem more adapted to intermittently connected networks such  
 128 as DTNs, yet a subset may allow CGrAnt to operate in different dynamic  
 129 scenarios.

### 130 *2.3. Related work*

131 We go through the related work in the area, discussing the most repre-  
 132 sentative results on both DTN forwarding protocol and swarm intelligence  
 133 methods.

#### 134 *2.3.1. DTN Forwarding protocols*

135 The most common solutions in the literature take a controlled flooding ap-  
 136 proach. For instance, epidemic routing provides an optimal solution in terms  
 137 of message delivery and latency, when no buffer constraint is present (Vahdat  
 138 and Becker, 2000). In Epidemic routing, a node buffers a message and passes  
 139 it on to all encountered nodes that have not received it before. No good  
 140 message forwarders prediction is performed. To limit resource utilization, a  
 141 hop-count field can be set in each message. Epidemic routing is simple and  
 142 provides high reliability and adaptability, but it might generate too many

143 redundant messages, wasting communication and battery resources. To re-  
 144 duce this overhead, the Spray and Wait approach (Spyropoulos et al., 2005)  
 145 sprays messages over different contacts and then wait for these contacts to  
 146 eventually deliver the message to the destination.

147 Predicted-based approaches try to reduce the message overhead by se-  
 148 lecting a few good relays. In this context and more related to CGrAnt,  
 149 several approaches estimate a delivery likelihood based on the frequency or  
 150 similarities of meeting with contacts like PROPHET (Lindgren et al., 2003),  
 151 Delegation Forwarding (Erramilli et al., 2008), and Spray and Focus (Spy-  
 152 ropoulos et al., 2007). In particular, in PROPHET, vectors are exchanged  
 153 that indicate the predictability of each node in delivering their messages. This  
 154 predictability increases every time two nodes come into contact and reduces  
 155 if they fail to meet frequently. When a node  $A$  establishes a contact with  
 156 a node  $B$ , a message will be sent to  $B$  if its message delivery’s prediction  
 157 is higher as compared to  $A$ . The delivery predictability also has a transi-  
 158 tive property. All these approaches, however, might be too conservative and  
 159 lose good forwarding opportunities in environments with scarce connectivity.  
 160 Most importantly, the majority of the approaches assume infinite buffers and  
 161 bandwidth.

162 Other approaches study the effect of social networking on data forward-  
 163 ing. BubbleRap (Hui et al., 2008) and SimBet (Daly and Haahr, 2007) use  
 164 information about social community structures and popularity within a com-  
 165 munity to choose good relays. (Zhang et al., 2012) introduce four social-aware  
 166 data diffusion schemes based on the social relationship and data similarity  
 167 of the contacts.

168 Differing from such protocols, CGrAnt conducts local and global searches  
 169 and gathers relevant information from the DTN nodes. CGrAnt can thus  
 170 analyze the utility of each node as a message forwarder and limit message  
 171 replications.

### 172 2.3.2. *Swarm intelligence methods*

173 Though ACO has been extensively used in network environments, espe-  
 174 cially in MANETs (Liu and Feng, 2005; Rosati et al., 2008), routing in DTNs  
 175 is challenging and few ACO protocols have been proposed. DAR (Rosati  
 176 et al., 2008) does not consider local information from neighboring nodes and  
 177 uses only the pheromone global information, which is not always available  
 178 in DTNs. ABMF (La and Ranjan, 2009) only aims to estimate the extra  
 179 capacity of each node as a message forwarder depending on its buffer dy-

namics. ACRP (Zhang et al., 2010b) uses the Epidemic protocol to flood the network with control messages associated with Forward Ants (FAs) and Backward Ants (BAs). None of the mentioned approaches consider the following important aspects of sparse and opportunistic networks: (i) analyzing social metrics of nodes, including their degree and betweenness centralities to aid select message forwarders; (ii) preventing the loss of previously found good paths caused by pheromone evaporation processes that are periodically performed (i.e., based on time, as in ABMF and ACRP) or the overuse of those paths due to the absence of an evaporation process (as in DAR); and (iii) dynamically limiting the number of control and data messages forwarded in the network.

(Ma et al., 2008; Zhang et al., 2010a) use CA for performing routing in a static topology with service quality constraints. They, however, operate in a static environment and do not analyze the dynamics of the contacts in a social network to determine opportunities. Moreover, they search for a single optimal path with a set of constraints and use Situational and Normative knowledge only to increase the convergence speed. Finally, they use a single and centralized belief space.

Considering these issues, we proposed a first version of the SI-based routing protocol for DTNs that used only ACO (Vendramin et al., 2012b). Guided by pheromone concentration, heuristic functions, and social metrics, the Greedy Ant (GrAnt) protocol performed better than the well-known DTN routing protocols (Vendramin et al., 2012b).

CGrAnt encompasses CA and ACO metaheuristics and can be considered as an extension of our previous method (Vendramin et al., 2012b). CGrAnt improves the learning process and the gathering, during evolution, of high-level information to be stored in the CA belief space.

### 3. The CGrAnt Routing Protocol

As previously mentioned, the CGrAnt routing protocol is based on CA and ACO meta-heuristics. CA is comprised of two spaces: (1) **Belief Space**, which represents the knowledge (i.e., set of information) gathered during the search for a set of paths and (2) **Population Space**, which is composed of individuals (i.e., ants) looking for solutions (i.e., forwarding paths) in an ACO framework.

In DTN, due to the lack of central element capable of storing and publishing all gathered information, the CGrAnt components are distributed among



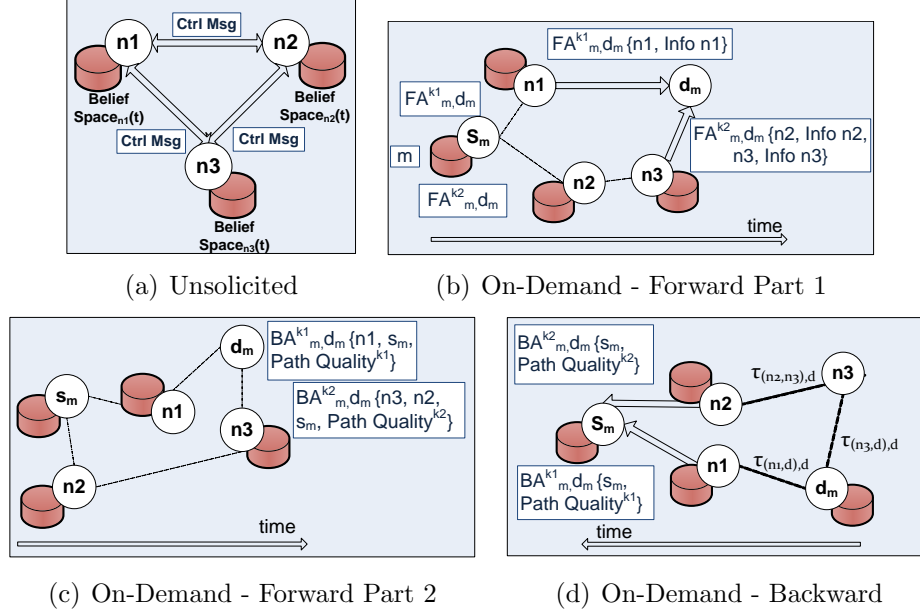


Figure 1: Operation Modes of the CGrAnt Protocol.

different belief spaces, each stored in a network node; each node knows only a subset of the population space. The exchange of information between the belief and population spaces always occurs in a distributed manner.

CGrAnt operates in two modes: **unsolicited** and **on-demand** (Figure 1). In the **unsolicited mode** (Figure 1(a)), control messages (Ctrl Msg) are always exchanged among neighboring nodes to update the information stored in each belief space. If it is necessary to establish a data session between the source of a data message  $m$  ( $s_m$ ) and its destination ( $d_m$ ), CGrAnt switches to the **on-demand mode** (Figures 1(b)- 1(d)). In each node that contains a data message  $m$  to be forwarded, one or more Forward Ants (FA)  $k$ , are forwarded toward  $d_m$  along with  $m$  via one or more neighboring nodes. During the path construction, an ant  $k$  collects information (Info) on each node  $n$  that composes the path toward  $d_m$ . The node  $n$  can be either a node  $i$  that contains a buffered message  $m$  to be forwarded or a neighboring node  $j$ . A subset of this information is used by CGrAnt for the belief space update of each node. The other part is carried by the FA until it reaches  $d_m$  (Figure 1(b)). In  $d_m$ , the quality of the constructed path is calculated based on the information gathered by the FA. A Backward Ant (BA) is sub-

sequently created with the information obtained by the corresponding FA (Figure 1(c)), the FA is deleted, and the BA is sent back through the reverse path followed by the FA. In its path toward the source ( $s_m$ ) of the FA, the BA updates the ACO pheromone concentration operator ( $\tau$ ) in each link between the nodes that compose the reverse path (Figure 1(d)) according to the constructed path quality. At each visited node, its identification is removed from the BA's record. In subsequent message forwarding, the ACO operators (Pheromone Concentration and Heuristic Function) and the CA belief space (along with other CGrAnt components) dictate the routing decision in each node and infer the best forwarders for each message.

The next sections describe the CGrAnt routing protocol in detail. Sections 3.1 to 3.3 describe the components used by CGrAnt that influence the search for paths. Sections 3.4 and 3.5 describe the routing phases of CGrAnt, which determine the route(s) a message must follow to reach its destination.

### 3.1. Metrics and Indicators

The communication between the belief and population spaces is mediated by specific metrics and indicators. The metrics incorporated into CGrAnt are classified as **basic** (obtained directly from the population space or nodes) or **composite** (obtained from manipulating basic metrics). The basic metrics are classified into **local** (associated with each node and its neighboring nodes) and **global** (associated with complete paths constructed by ants) categories. Figure 2 illustrates the metrics and their relationships with the belief space stored in each node. The Situational and History knowledge influence the population space and the population space update the global metrics.

Table 1 provides a brief description of the metrics and variables used throughout this paper.

The **Local Basic Metrics** used by CGrAnt include the following:

- **Frequency of Encounters** ( $FE_{n,d}$ ) between a pair of nodes  $n$  and  $d$ ;
- **Duration of an Encounter** ( $DE_{n,d}$ ) between  $n$  and  $d$ ;
- **Average Pause Time** ( $\overline{PT}_n$ ) in the places visited by  $n$ ;
- **Average Movement Speed** ( $\overline{MS}_n$ ) of a node  $n$ ;
- **Degree Centrality** ( $DC_n$ ) of a node  $n$  (Freeman, 1979). As  $n$  encounters more nodes in the network and increments its degree centrality, it has more opportunities to choose the best message forwarders;

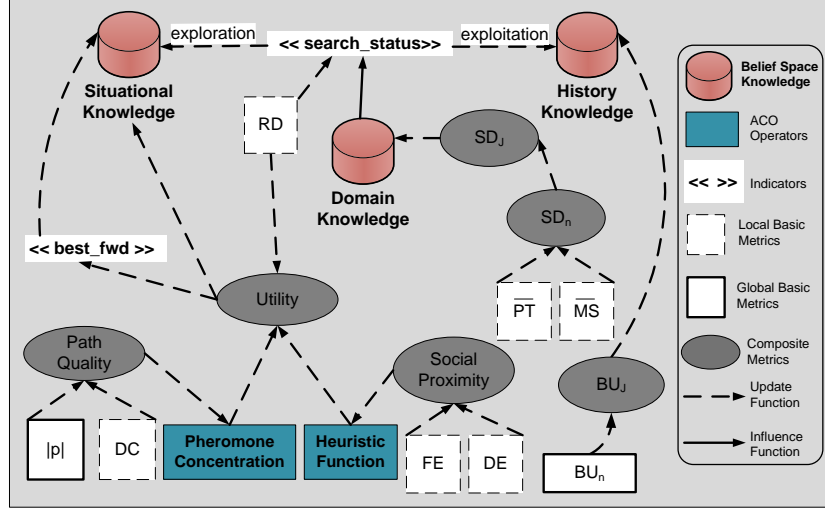


Figure 2: Belief Space of a Node.

- **Relationship Degree** ( $RD_{i,m}$ ) of a node  $i$  with respect to each of its data messages  $m$ .  $RD_{i,m}$  defines  $i$  as the source ( $s_m$ ) or an intermediate node of  $m$ .

The **Global Basic Metrics** of CGrAnt include:

- **Number of Hops** ( $|p|$ ) in a complete path  $p$ . If there are few hops in a path, fewer resources are consumed, and less interference is generated;
- **Betweenness Utility of a node** ( $BU_{n,d}$ )  $n$  relative to a destination node  $d$ . To obtain a high betweenness utility relative to  $d$ , a node  $n$  must appear with a high frequency in paths between any source node and  $d$ . Each time a node  $n$  receives a BA indicating that  $n$  is a component of a complete path (global solution) to  $d$ , its betweenness utility is updated,  $BU_{n,d}(t) = BU_{n,d}(t - 1) + 1$ .

The **Composite Metrics** of CGrAnt include the following:

- **Social Proximity** ( $SP_{n,d}$ ) between  $n$  and  $d$  is directly associated with the ACO local operator **Heuristic Function** ( $\eta$ ) and is defined as  $SP_{n,d} = FE_{n,d} \times DE_{n,d}$ ;

- 284 • **Path Quality** ( $Q_{p_{s_m, d_m}^k}$ ), which measures the quality of a path  $p$  con-  
 285 structed by an FA  $k$  between nodes  $s_m$  and  $d_m$ . It encompasses the  
 286 number of hops ( $|p|$ ) and the average degree centrality of nodes  $n$  ( $DC_n$ )  
 287 that compose a path to  $d_m$ :

$$Q_{p_{s_m, d_m}^k}(t) = \frac{\sum_{\forall n \in p_{s_m, d_m}^k} DC_n(t)}{|p|_{s_m, d_m}^k} + \frac{1}{|p|_{s_m, d_m}^k} \quad (1)$$

288 The Path Quality metric is directly associated with the ACO global  
 289 operator **Pheromone Concentration** ( $\tau$ );

- 290 • **Utility** of a node  $n$  in relation to  $d$  ( $U_{n,d}$ ), which describes how well  
 291  $n$  can perform as a message forwarder to  $d$ . The  $U_{n,d}$  is subsequently  
 292 determined according to the basic metric  $RD_{i,m}$ :

$$U_{n,d}(t) = \begin{cases} \eta_{n,d}(t) & \text{if } RD_{i,m} = s_m \\ \eta_{n,d}(t) + \tau_{(i,y),d}(t) & \text{otherwise,} \end{cases} \quad (2)$$

293 The  $U_{n,d}(t)$  can thus consider only local or both local and global in-  
 294 formation. The local information  $\eta_{n,d}(t)$  is the heuristic function of  
 295 ACO measured by  $SP_{n,d}$ , and the global information  $\tau_{(i,y),d}(t)$  is the  
 296 pheromone concentration on each link  $(i, y)$  belonging to a path to  $d$ ,  
 297 and  $y$  is defined as:

$$y = \begin{cases} d_m & \text{if } n = i \\ j & \text{if } n = j, \end{cases} \quad (3)$$

- 298 • **Stagnation Degree** ( $SD_n$ ) of a node  $n$ , which allows the identifica-  
 299 tion of the most mobile nodes in a dynamic scenario (e.g., buses or  
 300 vehicles) and consequently, adapts CGrAnt to heterogeneous network-  
 301 ing encounters on the fly. For this,  $SD_n$  considers the node average  
 302 pause time ( $\overline{PT}_n$ ) and average movement speed ( $\overline{MS}_n$ ):

$$SD_n(t) = \left( \frac{\overline{PT}_n}{\overline{MS}_n} \right) \quad (4)$$

- 303 • **Stagnation Degree of the social network** ( $SD_J^i$ ) of a node  $i$ , which  
 304 is based on the  $SD_n$  metric:

$$SD_J^i(t) = \frac{1}{|J|} \sum_{j \in J} SD_j(t), \quad (5)$$

305 where  $J$  is the set of nodes encountered by node  $i$ ;

- 306 • **Betweenness Utility of the social network** ( $BU_{J,d}^i$ ) of a node  $i$   
 307 in relation to  $d$ , which considers the basic metric  $BU_{n,d}$ . The  $BU_{J,d}^i$  is  
 308 initialized differently depending on the origin of the BA that announces  
 309 to  $i$  that a complete path to  $d$  has been constructed:

$$BU_{J,d}^i(0) = \begin{cases} BU_{i,d}(t) & \text{if BA came from } d \\ BU_{j,d}(t) & \text{if BA came from } j \end{cases} \quad (6)$$

310 The  $BU_{J,d}^i$  metric is updated with  $BU_{i,d}$  whenever  $i$  receives a BA.  
 311 When the BA comes from a neighboring node  $j$  (not from  $d$ ),  $BU_{J,d}^i$  is  
 312 also updated with  $BU_{j,d}$ :

$$BU_{J,d}^i(t) = \begin{cases} Z & \text{if BA came from } d \\ \frac{1}{2}(Z + BU_{j,d}(t)) & \text{if BA came from } j, \end{cases} \quad (7)$$

313 where  $Z = \frac{1}{2}(BU_{J,d}^i(t-1) + BU_{i,d}(t))$ .

314 More details on the metrics definitions are discussed in [23].

315 In addition to the basic and composite metrics, CGrAnt uses two indica-  
 316 tors: (1) **best\_fwd<sub>m</sub>**, which stores the current best forwarder for a specific  
 317 message  $m$  and (2) **search\_status**, which decides whether FAs must *explore*  
 318 or *exploit* the network while seeking solutions for the DTN forwarding prob-  
 319 lem.

### 320 3.2. Population Space

321 The population space is composed of FAs and BAs messages. The FAs  
 322 look for sets of possible paths, i.e., one set  $P_m$  for each pair  $(s_m, d_m)$ . As-  
 323 suming a total of  $M$  messages to be forwarded in the entire network, there  
 324 will be several paths sets  $\{P_1, \dots, P_m, \dots, P_M\}$  constructed simultaneously.  
 325 Every paths set  $P_m$  represents a group of solutions generated whenever a  
 326 message  $m$  originated in  $s_m$  must be sent to  $d_m$ . Each complete path  $p$  in  $P_m$   
 327 composed of a sequence of nodes is established when ant  $k$  reaches  $d_m$  and  
 328 can be defined as:

$$p_{s_m, d_m}^k = \{s_m, \dots, n, \dots, d_m\}, k = 1, \dots, K_m.$$

329 For each node  $i$  with a message  $m$ , each time a better forwarder  $j$  for  $m$   
 330 appears, a new FA  $k$  is generated to begin its path construction, and a copy  
 331 of  $m$  is sent to  $j$ . The partially constructed path from  $s_m$  to  $j$  remains the  
 332 same. However, from  $j$ , the FA  $k$  is free to find  $d_m$  passing through different  
 333 nodes  $n$  and can provide a different path into the population space. The  
 334  $K_m$  defines the number of FAs generated to find solutions for  $m$ . In general,  
 335  $|P_m| \ll K_m$  because a subset of ants cannot find  $d_m$ . The parameter  $K_m$   
 336 is dynamically defined according to the belief space knowledge and message  
 337 delivery success. Ant generation is thus auto-adaptive, as is the number of  
 338 constructed paths, and both depend on the DTN dynamics.

### 339 3.3. Belief Space

340 In CGrAnt protocol, there is no element to centralize and share the gath-  
 341 ered knowledge. The belief spaces are thus distributed over the network.  
 342 Each belief space in a node encompasses three types of knowledge: Domain,  
 343 History, and Situational, as detailed hereafter.

#### 344 3.3.1. Domain Knowledge

345 introduced to support the analysis of local and specific DTN dynamics.  
 346 Domain knowledge keeps CGrAnt updated on the relative local dynamics  
 347 of each node  $i$ , which is evaluated in this paper based on the relationship  
 348 between  $SD_i$  and  $SD_j^i$ . The knowledge is distributed among nodes,  $Dom_i$ ,  
 349 for  $i = 1, \dots, N$ , where  $N$  is the number of nodes available in the network  
 350 and  $Dom_i$  assists CGrAnt in characterizing  $i$  into three classes: a node with  
 351 high, medium or low stagnation degree.

352 Based on this information and the specific heuristics of a DTN forwarding  
 353 problem, the Domain knowledge can set the status of the path search (i.e.,  
 354 exploration or exploitation). The *Acceptance* and *Update* functions of the  
 355 Domain knowledge occur more frequently than its *Influence Function* because  
 356 they are called during the node encounters. The *Influence Function* acts only  
 357 during the message forwarding phase.

358 The *Acceptance Function* accepts information on the local dynamic of  
 359 each neighboring node  $j$ , only if  $SD_j \in (0, \infty)$ . After accepting a new so-  
 360 lution, the *Update Function* is called. It considers the event window  $W(t)$

361 containing a list of local events of  $i$ , such as an encounter between  $i$  and a  
 362 non-stationary node  $j$ .

363 The *Update Function* adds the pair  $(SD_j(t), j)$  to the list and updates  
 364  $SD_j^i$  according to Eq. 5.

365 The *Influence Function* evaluates the stagnation degree of  $i$  with respect  
 366 to its social network:

- 367 • *High stagnation degree*: A node  $i$  is characterized as a highly stagnated  
 368 node when the following relations apply:  $SD_i > SD_j^i + V_s$  and  $SD_j^i \in$   
 369  $(0, \infty)$ . In this case, its improvement stagnation direction  $dr_i$  is set to  
 370  $dr_i = +1$ ;
- 371 • *Low stagnation degree*: A node  $i$  has a low stagnation degree when the  
 372 following relations apply:  $SD_i < SD_j^i - V_i$  and  $SD_j^i \in (0, \infty)$  hold. In  
 373 this case,  $dr_i = -1$ ;
- 374 • *Medium stagnation degree*: A node  $i$  is characterized as a medium stag-  
 375 nation node when the previously described relations do not apply. In  
 376 this case,  $dr_i = 0$ .

377 The  $V_i$  and  $V_s$  respectively define the decrement and increment in  $SD_j^i$   
 378 used to define the range of nodes with medium stagnation degree.

379 Based on  $dr_i$ , the *Influence Function* changes the value of the *search\_status*  
 380 indicator, which aids in defining whether an FA in node  $i$  must be sent to  
 381 exploit or explore during the search for a solution of the DTN forwarding  
 382 problem.

### 383 3.3.2. **History Knowledge**

384 introduced to adapt CGrAnt to changes in the network, thus increasing  
 385 the ability to reflect the global network dynamics. This knowledge stores a  
 386 history of important past events (in this case, the information that complete  
 387 paths to  $d$  have been found). Due to the lack of a central component, the  
 388 complete paths cannot be stored in the History knowledge. The betweenness  
 389 utility metric (representing the partial information of a complete solution)  
 390 is subsequently used and may subsequently influence the path search. The  
 391 History Knowledge is distributed among the network,  $His_i$ , for  $i = 1, \dots, N$ .  
 392 In each node  $i$ , this knowledge is divided in a total of  $D_i$  sub-knowledge:  
 393  $His_i(t) = \{His_{i,1}(t), His_{i,2}(t), \dots, His_{i,D_i}(t)\}$ , where  $D_i \leq N$  represents  
 394 the number of destination nodes for which node  $i$  originated or intermediated  
 395 a path.

396 The *Acceptance and Update Functions* of the History knowledge are called  
 397 in the backward phase, and its *Influence Function* acts in the message for-  
 398 warding phase.

399 The *Acceptance Function* is called after node  $i$  receives a BA from a  
 400 neighboring node ( $j$  or  $d$ ), indicating that a complete path to  $d$  has just  
 401 been found. The betweenness utility of  $i$  with respect to  $d$  ( $BU_{i,d}$ ) is always  
 402 accepted to update the belief space of  $i$ . The betweenness utility of the  
 403 neighboring node  $j$  ( $BU_{j,d}$ ) is also accepted to update the belief space of  $i$   
 404 if and only if the BA did not come from  $d$ . After the acceptance phase, the  
 405 *Update Function* is called to calculate  $BU_{j,d}^i$ .

406 The  $BU_{j,d}$  and  $BU_{j,d}^i$  dynamics are considered by the *Influence Function*  
 407 when evaluating the improvement directions of each neighboring node  $j$  as  
 408 a candidate forwarder for message  $m$  to  $d$ :  $dr_j = +1$ , if  $BU_{j,d} > BU_{j,d}^i$ ;  
 409  $dr_j = -1$ , if  $BU_{j,d} < BU_{j,d}^i$ ; or  $dr_j = 0$ , if  $BU_{j,d} = BU_{j,d}^i$ . Based on these  
 410 directions and the *search\_status* indicator, the history knowledge influences  
 411 the message forwarding by deciding whether an FA should be sent through  
 412 a previously found path and thus exploit the path search.

### 413 3.3.3. *Situational Knowledge*

414 introduced in CGrAnt to provide a memory of the best solutions, and  
 415 to influence the search process for a set of paths through these solutions.  
 416 Its memory is partial instead of complete and is represented by the best  
 417 forwarder of each message. The Situational Knowledge is distributed among  
 418 the network nodes:  $Sit_i$ , for  $i = 1, \dots, N$ . In each node  $i$ , the Situational  
 419 knowledge is divided in a total of  $M_i$  sub-knowledge:

$$420 \quad Sit_i(t) = \{Sit_{i,1}(t), \dots, Sit_{i,m}(t), \dots, Sit_{i,M_i}(t)\},$$

421 where  $M_i$  represents the number of data message  $m$  stored in node  $i$ 's buffer.

422 The *Acceptance, Update, and Influence* functions of the Situational knowl-  
 423 edge are called during the message forwarding phase.

424 The *Acceptance Function* operates with the understanding that if a new  
 425 neighboring node  $j$  (partial solution) is found, it is accepted to update the  
 426 belief space of  $i$  only if  $U_{j,d} > U_{best\_fwd_m,d}$ , where  $U_{best\_fwd_m,d}$  is the current  
 427 best forwarder utility for  $m$  stored in the sub-knowledge  $Sit_{i,m}$ . The accep-  
 428 tance condition can be relaxed to accept nodes with the same utility of the  
 429 best forwarder if the corresponding *search\_status* exploration is true. Only  
 430 one solution is accepted in each update. After accepting a new solution, the  
 431 *Update and Influence* functions are called. The new solution thus replaces  
 432 the previous solution ( $best\_fwd_m = j$ ), and the new solution quality ( $U_{j,d}$ )



433 updates  $U_{best\_fwd_m,d}$  (i.e.,  $U_{best\_fwd_m,d} = U_{j,d}$ ). The *Influence* Function creates  
 434 a new FA and forwards it along with  $m$  to the new solution  $j$ . The *Influence*  
 435 function thus dictates the future of each path built for a pair  $s_m - d_m$  and  
 436 the number of generated ants.

437 In the following sections, we describe in more detail the two phases that  
 438 dictate the main functioning of the GrAnt routing protocol: (1) **Message**  
 439 **Forwarding** or **Path Search**, which is represented by FAs looking for a set  
 440 of paths and the updating of selected knowledge and metrics. In this phase  
 441 occurs the data message forwarding; and (2) **Backward**, which is represented  
 442 by BAs updating the knowledge and metrics stored in the nodes.

### 443 3.4. Message Forwarding Phase

444 The message forwarding phase in CGrAnt is initialized on-demand when  
 445 a message  $m$  stored in a node  $i$  must be delivered to  $d_m$ , as described by  
 446 Algorithm 1. The FAs are subsequently created, encapsulated into  $m$ , and  
 447 sent toward  $d_m$  via one or more neighboring nodes  $j$ .

448 During message forwarding, an FA at a node  $i$  decides whether to for-  
 449 ward  $m$  to a new neighbor  $j$  according to the influence of the three types of  
 450 knowledge stored in its belief space: Domain, History, and Situational.

451 Because node  $i$  is the first candidate solution for forwarding message  $m$ ,  
 452 its identification and utility initialize the Situational knowledge (lines 9 and  
 453 10). The decision on forwarding  $m$  to  $j$  may be to explore (i.e., it is not  
 454 required that  $j$  has previously participated in a path to  $d_m$ ) or exploit (i.e.,  $j$   
 455 has previously participated in a path toward  $d_m$ ). The decision is guided by  
 456 the information on  $i$  in terms of  $RD_{i,m}$  and  $SD_i$  stored in  $Dom_i$ . The status  
 457 is initialized, enabling both exploitation and exploration of the path search  
 458 space (lines 12 and 13). These conditions can change in two situations: (i)  
 459  $i$  is an intermediate node with a medium stagnation degree (in this case,  
 460 the exploration stops (line 16)), or (ii)  $i$  is an intermediate node with a low  
 461 stagnation degree (because  $i$  is a highly mobile node, it does not forward  $m$ ,  
 462 thus, both exploration and exploitation are stopped (lines 19 and 20)).

463 Aside from Domain knowledge, the History and Situational knowledge  
 464 also influence the CGrAnt message forwarding. The History knowledge in-  
 465 fluences the forwarding (line 33) when the decision is to exploit the solutions  
 466 already found:  $i$  forwards  $m$  to a solution  $j$  if node  $j$ 's betweenness util-  
 467 ity ( $BU_{j,d}$ ) is higher than the betweenness utility of node  $i$ 's social network  
 468 ( $BU_{J,d}^i$  stored in  $His_{i,d}$ , as observed in line 47). The Situational knowledge in-  
 469 fluences CGrAnt (line 44) when the decision is to explore the network or when

470 better solutions appear. Node  $i$  only forwards  $m$  to a new solution  $j$  if one of  
 471 the following two conditions is satisfied: (i) The utility of  $j$  ( $U_{j,d}$ ) is higher  
 472 than the utility of the best forwarder previously found for  $m$  ( $U_{best\_fwd_m,d}$ ,  
 473 which is stored in  $Sit_{i,m}$ , as observed in lines 53 and 54). This condition at-  
 474 tempts to locate a new best forwarder for  $m$  among the current neighboring  
 475 nodes of  $i$ ; (ii) At the beginning of the exploration (i.e.,  $m$  was not forwarded  
 476 to any other node and  $best\_fwd_m = i$  and  $U_{j,d} = U_{i,d}$ , as in lines 26-27, 36-  
 477 38). The Situational and History knowledge are important contributions of  
 478 CGrAnt because they dynamically control the number of created ants and  
 479 the data message redundancy by setting each new best message forwarder  
 480 or forwarding  $m$  to already known good nodes, thus differing from the pure  
 481 ACO algorithms proposed for DTNs.

482 After analyzing the utility of every current neighbor  $j$  and inferring the  
 483 best choice, CGrAnt sends  $m$  to the designated forwarder (lines 33 and/or  
 484 44).

485 In addition to the data message forwarding, control messages are period-  
 486 ically and locally exchanged between  $i$  and its neighboring nodes  $j$  to update  
 487  $Dom_i$ .

488 The search for new paths toward  $d_m$  continues until  $i$  performs one of  
 489 the following actions: (1) encounters  $d_m$ , (2) becomes aware of the successful  
 490 delivery of the corresponding data message to  $d_m$ , or (3) detects that the  
 491 Time to Live (TTL) field of the data message has expired.

492 Throughout its path search, an FA carries the following information: the  
 493 ID of  $s_m$ , the ID of  $d_m$ , the node IDs through which it passes (between  $s_m$  and  
 494  $d_m$ ), and the degree centrality of each visited node  $j$  ( $DC_j$ ). The individual  
 495 qualities update the partial quality of the path under construction by the FA.  
 496 When an FA (along with  $m$ ) reaches  $d_m$ , the final quality of the constructed  
 497 path ( $Q_{p_{s_m,d_m}^k}$ ) is calculated as in Eq. 1. After calculating the quality of each  
 498 new and complete path  $p$ , a new control message, the BA, is created from  
 499 the information obtained by the FA, and the FA is deleted.

### 500 3.5. Backward Phase

501 During the backward phase, the BA returns to the node that originated  
 502 the message  $m$  through the reverse path selected by the FA. The concept  
 503 of using a reverse path in DTNs is motivated by wireless social networks in  
 504 which (i) individuals are often linked by a short chain of acquaintances, (ii)  
 505 certain encounters show repetitive behavior, and (iii) nodes have routines  
 506 that result in frequently visited locations and encounters.

507 In the reverse path, receiving a BA sent from node  $y$  to each neighboring  
 508 node  $i$  produces three effects: (1) increasing the  $BU_{i,d}$  value by one; (2)  
 509 initializing or updating the  $BU_{j,d}^i$  value according to Eq. 6 and 7, an effect  
 510 that corresponds to the *Update Function* of the History Knowledge; and (3)  
 511 updating the pheromone concentration toward  $d$  according to:

$$\tau_{(i,y),d}(t) = (1 - \rho) \times \tau_{(i,y),d}(t - 1) + Q_{p_{s_m,d_m}^k}(t), \quad (8)$$

512 where  $\tau_{(i,y),d}(t - 1)$  is the pheromone on link  $(i, y)$  that was last updated  
 513 at time  $(t - 1)$ . The evaporation process  $(1 - \rho)$  is necessary for the ants "to  
 514 forget" the previous pheromone values deposited on a link to a specific  $d$ .  
 515 This evaporation reduces the influence of the path search history. When the  
 516 pheromone is updated, all concentrations that belong to  $d$  are evaporated in  
 517  $i$ .

518 Even if the BA does not reach the node that originated the message  
 519 (due to connection partitions), the message forwarding phase of the CGrAnt  
 520 protocol is guided by a local path search provided by the heuristic function  
 521 information. Differing from other protocols that use only global (pheromone  
 522 concentration) or local (heuristic function) information, CGrAnt contains ad-  
 523 ditional flexibility, because the decision is based on all available information  
 524 (both ACO operators and knowledge stored in the CA belief space).

525 Additionally, the BA serves as an acknowledgment that  $m$  has achieved  
 526  $d_m$ , allowing the nodes that still maintain  $m$  to delete it and its associated  
 527 variables. A node that encounters another node that has already received a  
 528 BA for a given data message also deletes the corresponding message and its  
 529 associated variables. When the source node receives it, the BA is deleted.  
 530 Full paths are thus constructed for each destination using the information  
 531 gathered by the ants during the path search phase.

## 532 4. Evaluation Methodology

533 This section describes the numerical analysis we conducted using the  
 534 Opportunistic Network Environment (ONE) Simulator (Keränen et al., 2010)  
 535 to investigate the benefits of the metrics and components incorporated into  
 536 our proposal. Using ONE we can also assess both performance and accuracy  
 537 of the CGrAnt protocol in simulation scenarios that consider two different  
 538 movement models: Working Day (WD) (Ekman et al., 2008) and Points of  
 539 Interest (PoI) (Keränen et al., 2010), both proposed by default in the ONE

540 simulator. The ONE default parameters were kept unchanged (whenever  
 541 possible) in order to confirm the results when compared with the others two  
 542 evaluated protocols (Epidemic and Prophet). Moreover, using the default  
 543 parameters, we intend to facilitate the dissemination of our proposal in the  
 544 simulation platform.

545 The WD movement model represents an activity-based environment that  
 546 simulates the daily lives (activities) of people who go to work in the morning,  
 547 spend the day working, may go to a public place for leisure activities with  
 548 friends at the end of day, and return to their houses at night. In WD, the  
 549 total area ( $10,000 \times 8,000 \text{ m}^2$ ) encompasses meeting points, buses, houses,  
 550 offices, and roads. The area is divided into four regions denoted by  $R_A$  to  
 551  $R_D$ . Eight groups of nodes, denoted by  $A$  to  $H$ , are created to represent the  
 552 node movements into specific regions. Groups  $A$  to  $D$  simulate only intra-  
 553 region movements (e.g., group  $A$  simulates the node movements into region  
 554  $R_A$ ). Groups  $E$ ,  $F$ , and  $G$  simulate node movements between  $R_A$  and other  
 555 regions, and  $H$  simulates movements among all regions. The assignment of  
 556 nodes per group is as follows:  $A$  has 258 nodes,  $B$  has 119,  $C$  has 154,  $D$  has  
 557 154,  $E$  has 102 (i.e.,  $E = A \cap B$ ),  $F$  has 122 (i.e.,  $F = A \cap C$ ),  $G$  has 122  
 558 (i.e.,  $G = A \cap D$ ), and  $H$  has 70 (i.e.,  $H = A \cap B \cap C \cap D$ ).

559 The WD movement model represents an activity-based environment that  
 560 simulates the daily lives (activities) of people who go to work in the morning,  
 561 spend the day working, may go to a public place for leisure activities with  
 562 friends at the end of day, and return to their houses at night. In WD, the  
 563 total area ( $10,000 \times 8,000 \text{ m}^2$ ) encompasses meeting points, buses, houses,  
 564 offices, and roads. The area is divided into four regions denoted by  $R_A$  to  
 565  $R_D$ . Eight groups of nodes, denoted by  $A$  to  $H$ , are created to represent the  
 566 node movements into specific regions. Groups  $A$  to  $D$  simulate only intra-  
 567 region movements (e.g., group  $A$  simulates the node movements into region  
 568  $R_A$ ). Groups  $E$ ,  $F$ , and  $G$  simulate node movements between  $R_A$  and other  
 569 regions, and  $H$  simulates movements among all regions. The assignment of  
 570 nodes per group is as follows:  $A$  has 258 nodes,  $B$  has 119,  $C$  has 154,  $D$  has  
 571 154,  $E$  has 102 (i.e.,  $E = A \cap B$ ),  $F$  has 122 (i.e.,  $F = A \cap C$ ),  $G$  has 122  
 572 (i.e.,  $G = A \cap D$ ), and  $H$  has 70 (i.e.,  $H = A \cap B \cap C \cap D$ ).

573 In the PoI movement model, the total area ( $8,800 \times 7,800 \text{ m}^2$ ) is divided  
 574 into five points of interest that simulate several communities of people who  
 575 eventually meet each other and exchange data. There are eight groups of  
 576 nodes (W1, X1, Y1, Z1, W2, X2, Y2, and Z2), each with different desti-  
 577 nation selection probabilities. The POIs movement model is similar to the

community model used by A. Lindgren et al. (Lindgren et al., 2003).

Groups W1, X1, Y1, and Z1 contain 30 nodes each. Groups W2, X2, Y2, and Z2 have five nodes. Every node has one home PoI that is more likely to be visited than the other PoIs; there is a high probability that nodes meet each other within their same home community and a low probability that they go to PoIs outside their home community. The nodes select a destination, move in this direction (with  $\overline{MS} \in [0.5, 1.5]$  m/s), wait during a pause time ( $\overline{PT}$ ) ranging from 100 to 200 seconds (for groups W1-Z1) or 4,000 to 5,000 seconds (for W2-Z2), and select the next destination, among other actions. Table 2 shows the settings and communication parameters applied in all experiments for the three protocols under comparison. Unless otherwise described, the parameters used in both scenarios are emphasized in Table 2.

In Section 5, we evaluate the CGrAnt performance with variations in the setting parameters. We analyze the CGrAnt performance in the PoI scenario in terms of the percentage of messages delivered to destinations. The setting parameters emphasized represent those that provide the best results for the CGrAnt protocol. Next, in Section 6, we investigate a subset of the CGrAnt components by evaluating the protocol performance for both scenarios (WD and PoI). Section 7 compares CGrAnt with the Epidemic and PROPHET protocols in both scenarios and considers different aspects of the communication network context. In all experiments, each message has a size of 500 KB representing its payload and includes an FA with 8 bytes representing its path quality. The BA size is 100 bytes on average (including the header, path quality, and path hops).

The results discussed in Sections 5, 6 and 7 are presented in terms of mean values and confidence intervals (at a 95% confidence level) for 30 runs in each scenario. Due to the normality characteristics of the data under consideration, we apply the ANOVA (ANalysis Of VAriance) parametric statistical test together with its post hoc follow-up analysis over the independent groups considered. The ANOVA statistical test returns a p-value  $> 0.05$  indicating (with 95% of confidence) that there are no statistical differences among the groups and a p-value  $< 0.05$  if there is at least one pair of groups with a statistically significant difference. The intervals shown in the graphs for the post hoc analysis are computed in such a way that (to a close approximation) the two configurations compared are significantly different if their intervals are disjoint and are not significantly different if their intervals overlap. In our case, the delivery ratio interval associated with each group is represented by

a horizontal line (with a circle representing its mean value). For each graph, we choose one group for emphasis (represented by a black horizontal line and delimited by two vertical dashed lines).

## 5. Setting the Metrics of CGrAnt

This section investigates how selected metrics and ACO operators can improve the communication among swarms in the population space and thus, assist in obtaining better solutions. Sections 5.1 and 5.2 analyze the influence of certain metrics associated with the ACO operators of CGrAnt. Section 5.3 analyzes the influence of selected metrics in characterizing the utility of each node (solution) as a message forwarder.

### 5.1. Heuristic Function

We first analyze different sources of information in the ACO local operator or Heuristic Function ( $\eta_{n,d}(t)$ ). Recall that a node  $n$  is selected from a set of candidates to forward a message  $m$  to its destination  $d$ . In this section, the CGrAnt performance is evaluated by considering each of the following metrics associated with  $\eta_{n,d}(t)$ : (1)  $SP_{n,d}$ , (2)  $DC_n$ , (3)  $BU_{n,d}$ , and (4)  $FB_n$ , which represents the free space available in the buffer of a node  $n$ .

The experimental results show that when CGrAnt uses the  $SP_{n,d}$  metric, it delivers more messages ( $58.93 \pm 0.19\%$ ) than the  $DC_n$  ( $47.21 \pm 0.17\%$ ), the  $BU_{n,d}$  ( $53.92 \pm 0.17\%$ ), or the  $FB_n$  ( $49.72 \pm 0.28\%$ ).

Figure 3 presents the *post hoc* analysis of ANOVA highlighting the  $SP_{n,d}$  metric. The  $SP_{n,d}$  metric guarantees the highest message delivery ratio. This associated to the fact that there are no overlaps among the intervals resulted from other metrics, lead us to conclude that  $SP_{n,d}$  provides better performance than the other metrics.

The main reasons for the better performance of the  $SP_{n,d}$  metric are listed as follows: (1)  $SP_{n,d}$  indicates the proximity of  $n$  relative to  $d$  because it provides an estimation of the probability of future encounters between  $n$  and  $d$ . Moreover,  $SP_{n,d}$  contains information available along all of the search process. (2)  $DC_n$  indicates the popularity of  $n$  relative to all other nodes in the network instead of specific information for the candidate forwarder and  $d$ , as in  $SP_{n,d}$ . (3)  $BU_{n,d}$  provides important information for the nodes that successfully intermediated a communication to  $d$ ; however, it is only available for  $n$  after a complete path has been found (which includes  $n$ ), and  $n$  has received the visit of a BA (as seen in Section 3.5). (4) In highly connected

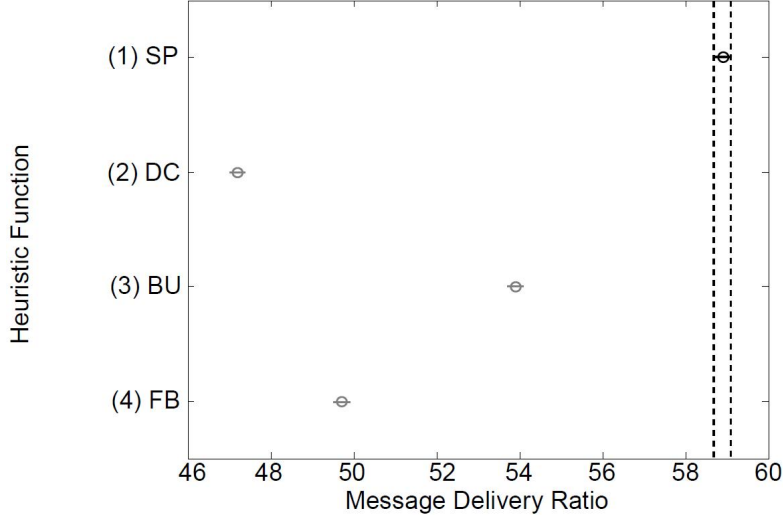


Figure 3: ANOVA results to Heuristic Function.

scenarios where the set of global solutions is always available and the same solution can be frequently selected, the percentage of available resources (e.g., the buffer) may be an important metric to consider; nevertheless, in an environment where the contacts are sparse (as in the DTNs), information on the social proximity between two nodes seems more important.

## 5.2. Pheromone Concentration

In this work, the pheromone concentration in the ACO global operator is associated with the quality of a complete path constructed by an ant  $k$  between the  $s_m$  and  $d_m$  nodes (i.e.,  $\tau_{(i,y),d}^k(t) = Q_{p_{s_m,d_m}^k}(t)$ ). Next, after defining the Heuristic function, we evaluate which type of information have more influence on  $Q_{p_{s_m,d_m}^k}(t)$ . The CGrAnt performance is evaluated using different metrics to predict the path quality: (1) the average Degree Centrality ( $DC$ ) of nodes  $n$  belonging to the path along with the reciprocal of the existing number of hops ( $|p|$ ) in the constructed path, i.e., as in Eq. 1; (2) only the second term of Eq. 1; (3) only the first term of Eq. 1; and (4) the average *Betweenness Utility* of node  $n$  relative to  $d$  (i.e.,  $Q_{p_{s_m,d_m}^k}(t) = \frac{\sum_n BU_{n,d}}{|p|}$ ).

The simulation results show that when using a composite metric encompassing  $DC$  and  $|p|$ , CGrAnt delivers more messages ( $58.93 \pm 0.19\%$ ) than when it uses only the basic metrics  $|p|$  ( $51.11 \pm 0.22\%$ ),  $DC$  ( $51.10 \pm 0.21\%$ ), and  $BU$  ( $49.89 \pm 0.22\%$ ).

Figure 4 presents the *post hoc* analysis of ANOVA highlighting the composite metric ( $DC$  and  $|p|$ ). As shown, the composite metric provides the best message delivery ratio. In addition, because there are no overlaps among the intervals of other metrics, we conclude that the composite metric provides better performance than the basic metrics. This result indicates that the node popularity is a good indicator of the node's ability to forward messages. This is particularly true in scenarios with intermittent connections, as in DTNs. Moreover, the importance of  $|p|$  is due to the fact that the smaller is the path, the fewer are the network resources consumed and the less is the communication interference that occurs.

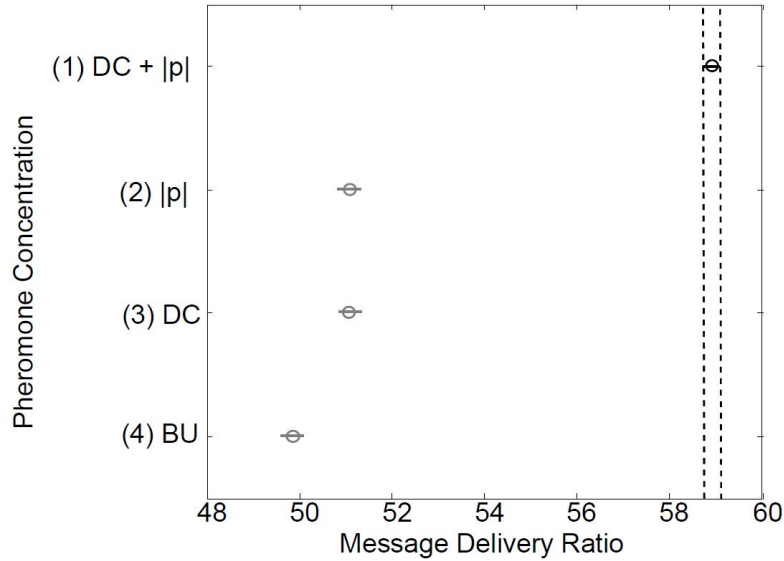


Figure 4: ANOVA results to Pheromone Concentration.

### 5.3. Node Utility

Finally, we analyze the utility ( $U_{n,d}$ ) of each node  $n$  as a forwarder of a message  $m$  to  $d$ . The CGrAnt performance is evaluated using four different metrics to describe the utility of  $n$  related to a reference node  $i$ , which contains a message  $m$  to be delivered to  $d$ . The investigated metrics are the following: (1) only local information represented by the *Heuristic Function* ( $U_{n,d} = \eta_{n,d}(t)$ ); (2) only global information represented by the *Pheromone Concentration* ( $U_{n,d} = \tau_{(i,y),d}(t)$ ); (3) the *Heuristic Function* and *Pheromone Concentration* ( $U_{n,d} = \eta_{n,d}(t) + \tau_{(i,y),d}(t)$ ); and (4) the *Heuristic Function*,



690 *Pheromone Concentration*, and the *Relationship Degree* ( $RD_{i,m}$ ) metric (ac-  
 691 cording to Eq. 2).

692 The results show that when both ACO operators (heuristic function and  
 693 pheromone concentration) and the  $RD_{i,m}$  metric are considered (composition  
 694 4), the CGrAnt protocol delivers more messages ( $58.93 \pm 0.19\%$ ) compared  
 695 with composition (1) in which it uses only the heuristic function ( $55.95 \pm$   
 696  $0.23\%$ ), composition (2) in which it uses only the pheromone concentration  
 697 ( $49.29 \pm 0.18\%$ ), and composition (3) in which it uses the heuristic function  
 698 and the pheromone concentration without the  $RD_{i,m}$  metric ( $58.02 \pm 0.20\%$ ).

699 Figure 5 presents the *post hoc* analysis of ANOVA highlighting the best  
 700 composition (4). As depicted, there is no overlap among the intervals. This  
 701 behavior verifies that the use of both local and global information (heuristic  
 702 function and pheromone concentration) along with the  $RD_{i,m}$  metric achieves  
 703 higher performance.

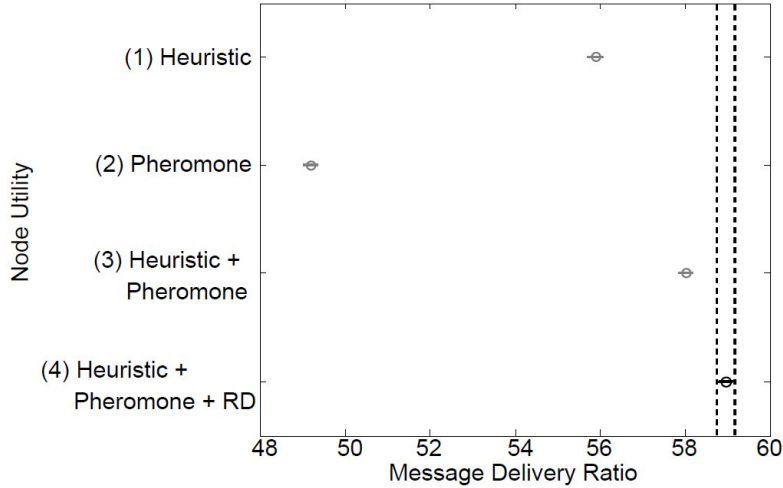


Figure 5: ANOVA results to Node's Utility Analysis.

## 704 6. CGrAnt Component Analysis

705 The influence of the CGrAnt's components on the message delivery ratio  
 706 and message redundancy ratio are evaluated for the PoI and WD scenarios.  
 707 The message redundancy ratio is expressed as  $Redundancy = (B_{transm} -$   
 708  $B_{delivery})/B_{delivery}$ , in which  $B_{transm}$  represents the number of bytes trans-

mitted to nodes, and  $B_{delivery}$ , which is the number of bytes delivered to their destination.

Initially, an additive methodology is adopted in which components are added one by one to the previous protocol configuration until its final version (configuration 6) is reached (see Table 3). The first two configurations represent the CGrAnt protocol that considers only the ACO metaheuristic. The pheromone concentration slightly increases the delivery ratio and reduces the number of replicated messages. The influence of the new components incorporated into CGrAnt (i.e., CA's belief space) is analyzed for configurations 3 to 6.

When comparing the results obtained from the pure ACO metaheuristic (configuration 2) with the final performance of CGrAnt (configuration 6), the following gains are achieved:

- in the POI scenario, the configuration 2 obtains 48.61% against 58.93% of message delivery in the configuration 6. This represents a percentage increase of message delivery in 21.23%;
- in the WD scenario, the configuration 2 obtains 54.97% against 63.19% of message delivery in the configuration 6. This represents a percentage increase of message delivery in 14.95%;
- in the POI scenario, the configuration 2 obtains 15.97% against 10.37% of message redundancy in the configuration 6. This represents a percentage decrease of message redundancy in 35.07%;
- in the WD scenario, the configuration 2 obtains 68.62% against 12.43% of message delivery in the configuration 6. This represents a percentage decrease of message redundancy in 81.86%.

The Situational knowledge (configuration 4) aims to dynamically restrict the number of FAs (and, consequently, the number of messages replicated) to only the most promising forwarders. The History knowledge (configuration 5) provides the exploitation of already known good solutions and consequently increases the message delivery ratio. The Domain knowledge 1 (configuration 3) privileges the exploration of the search space and the Domain knowledge 2 (configuration 6) favors the exploitation. According to the analysis of configurations 3 and 6, the Domain knowledge aims to increase the message delivery and reduce the redundancy ratio.

GrAnt, our previous ACO-based protocol presented in (Vendramin et al., 2012b), performs better than CGrAnt in PoI (i.e., GrAnt provides  $62.10 \pm 0.2\%$  of message delivery and  $7.05 \pm 0.1$  of message redundancy). However, for the WD scenario, CGrAnt outperforms GrAnt (the latter achieved  $60.25 \pm 0.75\%$  for delivery and  $13.65 \pm 0.14$  for redundancy). CGrAnt presents better performance in WD rather than in POI scenario because of the Domain Knowledge. The Domain Knowledge considers the node mobility in order to determinate if a node is a good message forwarder and, consequently, if that node must exploit or explore the search space. In WD there are several mobility patterns, i.e., buses and vagabonds nodes with high mobility. Otherwise, in POI, the mobility of nodes is very similar and the information of the Domain knowledge is less relevant.

Additionally, CGrAnt has the advantage of modeling at a higher abstraction level that enables the elimination of any knowledge of the CA belief space in a simple way (e.g., the history knowledge can be eliminated when the environment is more connected and fewer messages need to be forwarded).

Table 4 represents the eliminatory analysis of the CA's belief space proposed by CGrAnt. The first configuration in Table 4 represents the CGrAnt protocol including all components. In analyzing this table we conclude the following:

- **Domain Knowledge** aims to increase the message delivery ratio and reduces the message replication. This is particularly true in the WD scenario in which the nodes generally have a stagnation degree ( $SD_n$ ) lower than the average stagnation of its social network ( $SD_J$ ), and consequently, the Domain knowledge has greater influence. Without this knowledge, the message delivery ratio is reduced by 1.77% (PoI) and 1.40% (WD) and the message redundancy ratio increased by 0.77% (PoI) and 92.27% (WD);
- **Situational Knowledge** aims to dynamically restrict the FAs to only good forwarders, and its influence on the message redundancy ratio is therefore greater. Without this knowledge influence, we observe an increase of 51.78% (PoI) and 248.51% (WD) in the message redundancy ratio and a reduction in the message delivery ratio of 9.15% (PoI). In the WD scenario we have an increase of message delivery ratio of 1.65%;
- **History Knowledge** aims to enhance already known good solutions and thus increases the message delivery ratio. Without the influence of

779 this knowledge, there is a reduction of 3.78% (PoI) and 3.4% (WD) in  
780 the message delivery ratio with a lowest cost incurred in the replication  
781 of messages -33.85% (PoI) and -12.55% (WD).

782 It may be noted that the use of the multiple knowledge incorporated in  
783 CGrAnt improves its final performance. As the use of multiple techniques in  
784 ensemble systems, the combination of a set of techniques on an suitable con-  
785 sensus function provides better performance than each individual technique.

## 786 7. The CGrAnt Overall Performance

787 This section investigates how CGrAnt performs as a forwarding protocol  
788 when compared with the Epidemic and PROPHET protocols under varying  
789 networking parameters. We performed 30 runs, and the reported results rep-  
790 resent the mean and confidence intervals (at a 95% confidence level) values.  
791 To evaluate the reliability and the cost of the three protocols, we consid-  
792 ered the following three performance metrics: (1) message delivery ratio, (2)  
793 message redundancy ratio, and (3) average message delivery delay.

### 794 7.1. Analysis of Different Buffer Sizes

795 Figure 6 depicts the performance of the three protocols with variation of  
796 the buffer sizes (from 4 MB to 16 MB) for the PoI scenario. The dashed  
797 curves with empty points denotes the results for nodes that operate at a  
798 communication range of 10 m and a transmission rate of 2 Mbps (repre-  
799 senting bluetooth devices). The solid curves with black points show the  
800 results for nodes that operate at a 100 m range and a transmission rate of  
801 10 Mbps (representing WiFi devices). Figures 6(a) and 6(b) show that with  
802 the use of CGrAnt, more messages are delivered and less buffer space is de-  
803 voted to message replications. For instance, for a buffer size of 8MB and a  
804 communication range of 100 m, CGrAnt delivers  $93.27 \pm 0.10\%$  of messages  
805 (versus  $66.52 \pm 0.16\%$  for Epidemic and  $46.84 \pm 0.15\%$  for PROPHET) with  
806 a message replication of only  $15.23 \pm 0.06\%$  ( $38.90 \pm 0.13\%$  for Epidemic and  
807  $41.85 \pm 0.19\%$  for PROPHET). Figure 6(c) shows that CGrAnt provides the  
808 lowest delivery delay when using a higher buffer size (i.e., 10 MB to 16 MB).

809 Note that for buffer sizes lower than 8MB, PROPHET presents better  
810 results in terms of delivery delay. This is due to its lowest Message Delivery  
811 ratio, almost -50% than CGrAnt, only short route with short delivery delay  
812 is used.

### 813 7.2. Analysis of Different Message TTLs

814 Figure 7 shows the performance of the CGrAnt, Epidemic, and PROPHEt  
815 protocols with variation of the message TTL (Time-To-Live, i.e., how long  
816 the message lives in the network in minutes) for the PoI scenario. Figures 7(a)  
817 and 7(b) show that CGrAnt provides the best results in terms of message  
818 delivery and redundancy ratios for all message TTLs and both communica-  
819 tion ranges. For instance, with a 10 m range and a TTL of 2,100 minutes,  
820 CGrAnt delivers  $61.09 \pm 0.23\%$  of messages (versus  $35.39 \pm 0.17\%$  for Epi-  
821 demic,  $30.50 \pm 0.14\%$  for PROPHEt), with a message redundancy of only  
822  $10.21 \pm 0.05\%$  ( $30.47 \pm 0.18\%$  for Epidemic,  $32.82 \pm 0.17\%$  for PROPHEt).  
823 These results show that a node with an efficient routing protocol such as  
824 CGrAnt, with guidance from the CA knowledge and the ACO operators is  
825 able to efficiently manage message forwarding and dynamically limit message  
826 redundancy. Figure 7(c) shows that PROPHEt provided the best results in  
827 terms of delivery delay; this is the only metric for which CGrAnt cannot pro-  
828 vide the best results, a lack that is justified by its lowest Message Delivery  
829 ratio, almost  $-40\%$  than CGrAnt, only short route with short delivery delay  
830 is used.

831 The performance of the three DTN protocols with variations in the buffer  
832 sizes and message TTLs in the WD scenario is presented in (Vendramin et al.,  
833 2012a). The results in (Vendramin et al., 2012a) show that CGrAnt achieves  
834 a higher message delivery ratio and a lower redundancy ratio than those of  
835 Epidemic and PROPHEt.

### 836 7.3. Analysis of Different Simulation Times

837 We also perform an experiment to evaluate the number of messages de-  
838 livered by the three protocols along the simulation time in the PoI (4MB of  
839 buffer size and message TTL of 600) and the WD scenarios (10MB of buffer  
840 size and message TTL of 1,800), both with a communication range of 10m.  
841 The aim in this section is to demonstrate that a better delivery ratio can be  
842 achieved as the time increases.

843 Figure 8 shows the message delivery ratio obtained by the three protocols  
844 over time in the PoI and WD scenarios with a 10 m range. In CGrAnt, the  
845 performance gain is greater mainly in the WD scenario. When the simulation  
846 time is increased from 400,000 to 2,800,000 seconds, CGrAnt delivers slightly  
847 better performance via an increase of 7.37% (PoI) and 17.45% (WD) in the  
848 message delivery ratio. This gain is justified by the fact that the more is the  
849 gathered information by CGrAnt over time, the better are the choices it can

850 make concerning message forwarding candidates. In contrast, PROPHET  
851 and Epidemic did not exhibit any significant variation of performance when  
852 the simulation time was increased.

#### 853 7.4. Analysis of Operation Costs

854 Another important consideration in protocol performance is the cost of  
855 initializing/updating and storing the state of the network and nodes. This  
856 cost covers the amount of the following types of information: (1) locally  
857 exchanged between every two nodes  $i$  and  $j$  during any contact opportunity,  
858 and (2) locally stored in each node  $i$  of the network. For this analysis, we  
859 considered the PoI scenario with the emphasized parameters presented in  
860 Table 2.

861 For the storage cost in terms of the total number of bytes exchanged be-  
862 tween nodes  $i$  and  $j$  during every contact opportunity, PROPHET displays a  
863 higher cost due to the exchange of its delivery predictability list. Due to the  
864 transitive property of PROPHET, the number of records in the predictabil-  
865 ity list of a node rapidly reaches the total number of network nodes (139,  
866 excluding itself). Thus, the number of bytes exchanged in both directions  
867 of the contact is 2,224 bytes ( $139 \times 16$  bytes), as it is shown in Table 5.  
868 For the CGrAnt protocol, during the message forwarding phase, 16 bytes  
869 are sent by  $i$  to its neighboring node  $j$ , identifying a data message  $m$  to be  
870 sent and the stagnation degree of  $i$ . At the same time, 29 bytes are sent  
871 by  $j$  to  $i$  representing node  $j$  on how well it can perform as a forwarder for  
872  $m$ : including its stagnation degree, its degree centrality, its social proximity  
873 with the destination  $d$  of the message  $m$  stored in node  $i$ , its betweenness  
874 utility relative to  $d$ , and an indication (true or false) that it knows that  $m$   
875 was already received by  $d$ .

876 For the cost in terms of the total number of bytes stored in each node  $i$ ,  
877 because Epidemic relies on the message replications to eventually deliver its  
878 messages, its storage cost is null (i.e., it is a stateless protocol). Although  
879 CGrAnt generates a higher storage cost compared with PROPHET and Epi-  
880 demic, in the worst case, that cost (11.28 KB) represents only 0.28% of the  
881 total capacity of a node buffer, if considering a limiting buffer size of 4 MB  
882 per node.

883 Finally, we investigate the operational cost provided by CGrAnt when  
884 considering only its control messages related to the ACO ants (FAs and  
885 BAs) in both PoI (*buffer* of 4 MB and message TTL of 600 minutes) and WD  
886 (*buffer* of 10 MB and message TTL of 1,800 minutes) scenarios. In the PoI

887 scenario, the control bytes corresponding to the FAs and BAs represent only  
 888  $0.0131 \pm 0,0009\%$  (communication range of 10 m) and  $0.0186 \pm 0,0002\%$  (100  
 889 m range) of the total bytes generated by CGrAnt (counting the bytes related  
 890 to data messages, FAs, and BAs). In the WD scenario, the FAs and BAs  
 891 bytes represent, respectively,  $0.0194 \pm 0,0003\%$  (10 m) and  $0.079 \pm 0,0022\%$   
 892 (100 m) of the total bytes generated. Nevertheless, even if accounting for the  
 893 extra cost of these control bytes (for the FAs and BAs) in the total amount (in  
 894 bytes) of replicated messages in the network, CGrAnt propagates fewer bytes  
 895 in the network compared with Epidemic and PROPHEt due to the high  
 896 number of data messages replicated by the latter protocols. When compared  
 897 with Epidemic, CGrAnt provides a reduction of  $39.99 \pm 0,16\%$  (in PoI with  
 898 a 10 m range),  $59.60 \pm 0,11\%$  (PoI with a 100 m range),  $83.60 \pm 0,19\%$   
 899 (WD with a 10 m range), and  $89.49 \pm 0,26\%$  (WD with a 100 m range)  
 900 in the total number of bytes generated in the network. Similarly, when  
 901 compared with PROPHEt, these reductions are  $34.86 \pm 0,18\%$  (PoI with 10  
 902 m),  $43.37 \pm 0,15\%$  (PoI with 100 m),  $74.59 \pm 0,34\%$  (WD with 10 m), and  
 903  $78.57 \pm 0,56\%$  (WD with 100 m) in total bytes generated in the network.  
 904 Therefore, we conclude that with the smallest generated overhead, CGrAnt  
 905 is able to choose the best message forwarders and reduce the total number  
 906 of data bytes replicated in the network.

907 It is important to highlight that the algorithmic complexity for the CGrAnt  
 908 protocol is linear  $[O(n)]$  in the number of network nodes, as shown in Algo-  
 909 rithm 1.

## 910 8. Conclusions

911 The importance of inferring the social behavior of nodes to efficiently  
 912 deliver data in mobile and intermittently connected networks has motivated  
 913 the development of the hybrid swarm intelligence-based CGrAnt protocol.  
 914 Using a greedy version of ACO and CA, CGrAnt characterizes the utility of  
 915 each node as a message forwarder by considering a set of social-aware met-  
 916 rics. We performed a set of experiments to analyze the influence of selected  
 917 metrics associated with the ACO operators and the use of different metrics to  
 918 characterize the utility of each node as a message forwarder. Once the group  
 919 of metrics was set, we analyzed the influence of the ACO operators and CA  
 920 knowledge on the CGrAnt performance. Finally, we compared the perfor-  
 921 mance of CGrAnt with the Epidemic and PROPHEt protocols under varying  
 922 networking parameters. The simulation results showed that CGrAnt outper-

923 formed PROPHET and Epidemic forwarding protocols in terms of message  
 924 delivery (gains of 99.12% compared with PROPHET and 40.21% compared  
 925 with Epidemic) and message replication (63.60% lower than PROPHET and  
 926 60.84% lower than Epidemic). In addition, despite a higher storage cost com-  
 927 pared to PROPHET and Epidemic (11.28 KB in the worst case), CGrAnt  
 928 propagates fewer bytes in the network due to the high number of data mes-  
 929 sages replicated by the latter protocols (a reduction of 43.37% when compared  
 930 to PROPHET and 59.60% when compared to Epidemic). In future work, we  
 931 intend to study in more details the adaptive capabilities of CGrAnt, when  
 932 operating in a scenario with varying mobility conditions, i.e., from an almost  
 933 static to a completely mobile and disconnected networking environment. For  
 934 this work, we will investigate the self-adaptation of social-aware metrics,  
 935 which can be combined with and applied to each CGrAnt component. Fi-  
 936 nally, the comparison of CGrAnt performance with other related social-based  
 937 forwarding protocols as well as the use of real data sets is such analysis will  
 938 be also let for future works.

## 939 9. Curriculum Vitae

940 **Ana Cristina B. Kochem Vendramin** received her graduate degree  
 941 in Computer Science from Paranaense University in 2000 and Ph.D. degree  
 942 in Computer Engineering from UTFPR in 2012. Currently, she is a Professor  
 943 in DAINF Department at UTFPR. Her research interests include distributed  
 944 systems, swarm intelligence, routing in cellular and mobile ad hoc networks.

945 **Anelise Munaretto** received her Dipl.-Ing. degree in Computer En-  
 946 gineering from PUPCPR in 1994 and Ph.D. degree in Computer Networks  
 947 from the University of Paris VI in 2004. Currently, she is Associate Profes-  
 948 sor at UTFPR. Her research interests include routing and quality of service  
 949 optimization in wireless networks.

950 **Myriam Regattieri Delgado** received her graduate degree at Electrical  
 951 Engineering from UFG in 1990 and Ph.D. degree in Computer Engineering  
 952 from UNICAMP in 2002. Currently, she is associate professor at UTFPR.  
 953 Her research interests include natural computing, swarm intelligence applied  
 954 in computer networks.

955 **Aline Carneiro Viana** is a Researcher at INRIA Saclay. She got her  
 956 Ph.D. in Computer Science from the University Pierre et Marie Curie in 2005.  
 957 After holding a postdoctoral position at IRISA/INRIA Rennes, she joined



958 INRIA Saclay in 2006. Her research is primarily in success data delivery in  
959 wireless self-organized and adaptive networks.

960 **Mauro Fonseca** received the B.Sc Degree in Computers Engineering  
961 from the PUCPR in 1994 the Ph.D. degree in Computer Networks from  
962 University Pierre et Marie Curie in 2003. He is currently an associate profes-  
963 sor at UTFPR. His research interests include wireless networks and network  
964 performance.

## 965 Acknowledgements

966 This work is partially supported (1) by the Brazilian National Research  
967 Council (CNPq), under research grants 309571/2014-6 to A. Munaretto and  
968 309197/2014-7 to M.R. Delgado, and 479159/2013-0 to A. Munaretto and  
969 (2) by the STIC AmSud UCOOL project.

## 970 References

- 971 Cerf, V., Burleigh, S., Hooke, A., Torgerson, L., Durst, R., Scott, K.,  
972 Fall, K., Weiss, H., 2007. Delay-Tolerant Network Architecture. IETF  
973 RFC-4839, DTN Research Group.  
974 URL <http://www.ietf.org/rfc/rfc4838.txt> (Last access:  
975 March, 2015)
- 976 Chaintreau, A., Hui, P., Scott, J., Gass, R., Crowcroft, J., Diot, C., 2007. An  
977 improved ant-based algorithm for the degree-constrained minimum span-  
978 ning tree problem. *IEEE Trans. Mobile Computing* 6 (6), 606–620.
- 979 Daly, E. M., Haahr, M., 2007. Social network analysis for routing in dis-  
980 connected delay-tolerant MANETs. In: *Proceedings of the 8th ACM In-*  
981 *ternational Symposium on Mobile Ad Hoc Networking and Computing.*  
982 *Montreal, Canada*, pp. 32–40.
- 983 Dorigo, M., Gambardella, L. M., 1997. Ant colony system: A cooperative  
984 learning approach to the traveling salesman problem. *IEEE Trans. Evol.*  
985 *Comput.* 1 (1), 53–66.
- 986 Dorigo, M., Maniezzo, V., Colorni, A., 1996. The ant system: Optimization  
987 by a colony of cooperating agents. *IEEE Trans. Syst., Man, Cybern. B*  
988 26 (1), 29–41.

- 989 Ekman, F., Keränen, A., Karvo, J., Ott, J., 2008. Working day movement  
990 model. In: Proceedings of 1st ACM/SIGMOBILE Workshop on Mobility  
991 Models. Hong Kong, China, pp. 33–40.
- 992 Erramilli, V., Crovella, M., Chaintreau, A., Diot, C., 2008. Delegation for-  
993 warding. In: Proceeding of Mobihoc 2008. Hong Kong, China, pp. 251–259.
- 994 Freeman, L. C., 1979. Centrality in social networks: Conceptual clarification.  
995 *Social Networks* 1 (3), 215–239.
- 996 Gonzalez, M. C., Hidalgo, C. A., Barabasi, A.-L., 2008. Understanding indi-  
997 vidual human mobility patterns. *Nature* 453 (5), 779–782.
- 998 Hui, P., Crowcroft, J., Yoneki, E., 2008. BUBBLE Rap: Social-based for-  
999 warding in delay tolerant networks. In: Proceedings of Mobihoc 2008. Hong  
1000 Kong, China, pp. 241–250.
- 1001 Keränen, A., Kärkkäinen, T., Ott, J., 2010. Simulating mobility and DTNs  
1002 with the ONE. *J. of Communications* 5 (2), 92–105.
- 1003 Khabbaz, M. J., Assi, C. M., Fawaz, W. F., 2012. Disruption-tolerant net-  
1004 working: A comprehensive survey on recent developments and persisting  
1005 challenges. *IEEE Commun. Surveys Tuts.* 14 (2), 607–640.
- 1006 La, R. J., Ranjan, P., 2009. Ant-based adaptive message forwarding scheme  
1007 for challenged networks with sparse connectivity. In: Proceedings of the  
1008 28th IEEE Conference on Military Communications (MILCOM). Boston,  
1009 USA, pp. 1–7.
- 1010 Lindgren, A., Doria, A., Schelén, O., 2003. Probabilistic routing in intermit-  
1011 tently connected networks. *SIGMOBILE Mob. Comput. Commun. Rev.*  
1012 7 (3), 19–20.
- 1013 Liu, L., Feng, G., 2005. Swarm intelligence based node-disjoint multi-path  
1014 routing protocol for mobile ad hoc networks. In: Proceedings of the Fifth  
1015 International Conference on Information, Communications and Signal Pro-  
1016 cessing (ICICS). Bangkok, Thailand, pp. 598–602.
- 1017 Ma, J., p. Zhang, J., Yang, J., l. Cheng, L., 2008. Research on cultural  
1018 algorithm for solving routing problem of mobile agents. *Journal of Chine*  
1019 *Univ. of Posts and Telecom.* 15 (4), 121–125.

- 1020 Reynolds, R. G., 1994. An introduction to cultural algorithm. In: Proceed-  
1021 ings of the 3rd Annual Conference on Evolutionary Programming. Vol. 41.  
1022 San Diego, USA, pp. 131–139.
- 1023 Rosati, L., Berioli, M., Reali, G., 2008. On ant routing algorithms in ad hoc  
1024 networks with critical connectivity. *Ad Hoc Networks* 6 (6), 827–859.
- 1025 Spyropoulos, T., Psounis, K., Raghavendra, C. S., 2005. Spray and wait:  
1026 An efficient routing scheme for intermittently connected mobile networks.  
1027 In: Proceedings of the 2005 ACM SIGCOMM Workshop on Delay-tolerant  
1028 Networking. WDTN '05. Philadelphia, USA, pp. 252–259.
- 1029 Spyropoulos, T., Psounis, K., Raghavendra, C. S., 2007. Spray and focus: Ef-  
1030 ficient mobility-assisted routing for heterogeneous and correlated mobility.  
1031 In: Proceedings of Fifth Annual IEEE International Conference on Per-  
1032 vasive Computing and Communications Workshops. IEEE PerCom Work-  
1033 shops '07. White Plains, NY, USA, pp. 79–85.
- 1034 Tournoux, P.-U., Leguay, J., Benbadis, F., Whitbeck, J., C., V., de Amorim,  
1035 M. D., 2011. Density-aware routing in highly dynamic dtns: The rollernet  
1036 case. *IEEE Trans. Mobile Computing* 10 (12), 1755–1768.
- 1037 Vahdat, A., Becker, D., 2000. Epidemic routing for partially connected ad  
1038 hoc networks. Tech. rep., duke University. Technical Report-CS-2000-06.  
1039 URL <http://www.cs.duke.edu/~vahdat/ps/epidemic.pdf> (Last  
1040 access: April, 2015)
- 1041 Vendramin, A. C. B. K., Munaretto, A., Delgado, M., Viana, A. C., 2012a.  
1042 CGrAnt: A swarm intelligence-based routing protocol for delay tolerant  
1043 networks. In: Proceedings of the 14th Annual Conference on Genetic and  
1044 Evolutionary Computation. GECCO '12. Philadelphia, USA, pp. 33–40.
- 1045 Vendramin, A. C. B. K., Munaretto, A., Delgado, M., Viana, A. C.,  
1046 2012b. GrAnt: Inferring best forwarders from complex networks' dynam-  
1047 ics through a greedy ant colony optimization. *Computer Networks* 56 (3),  
1048 997–1015.
- 1049 Zhang, M.-W., Sun, X.-M., Lv, X.-Y., 2010a. A QoS routing algorithm based  
1050 on culture-ant colony algorithm. In: Proceedings of the International Con-  
1051 ference on Computer Application and System Modeling. Taiyuan, China,  
1052 pp. 198–201.

- 1053 Zhang, P., Wang, H., Xia, C., Lv, L., Liu, X., 2010b. ACRP: Ant-colony-  
1054 based routing protocol for DTMNs. In: Proceedings of the Interna-  
1055 tional Conference on Educational and Information Technology. Chongqing,  
1056 China, pp. 272–276.
- 1057 Zhang, Y., Gao, W., Cao, G., Porta, T. L., Krishnamachari, B., Iyengar, A.,  
1058 2012. Social-aware data diffusion in delay tolerant manets. Handbook of  
1059 Optimization in Complex Networks. Springer Optimization and Its Appli-  
1060 cations 58, 457–481.

Table 1: Summary of the metrics and variables used to describe the network and CGrAnt

Network Variables	
$m$	A specific data message
$s$	A general source node ( $s_m$ : refers to the source of $m$ )
$d$	A general destination node ( $d_m$ : the destination node of $m$ )
$i$	Node with a data message to be forwarded
$j$	Neighboring node of $i$
$n$	A generic node, $i$ or $j$
$y$	A generic node, $j$ or $d$ in a link toward $d$
$N$	The total number of nodes available in the network
$M$	The total number of data messages to be forwarded
$J$	Set of nodes $j$ encountered by the node $i$ (social network of $i$ )
$D_i$	The total number of destination nodes for which $i$ originated or intermediated a path
CGrAnt Variables	
$FA/BA$	A general Forward/Backward Ant
$k$	A specific FA/BA
$K_m$	The total number of FAs generated for a message $m$
$p$	A complete path with a total of $ p $ hops
$P_m$	Group of paths $p$ constructed by ants for a message $m$
$W$	Event window
$dr_n$	Improvement direction of a node $n$
$V_i/V_s$	Inferior/Superior limits which define the range of medium stagnation degree
CGrAnt Local Metrics	
$FE_{n,d}$	Frequency of Encounters between $n$ and $d$
$DE_{n,d}$	Duration of an Encounter between $n$ and $d$
$PT_n$	Average Pause Time in the places visited by a node $n$
$MS_n$	Average Movement Speed of a node $n$
$DC_n$	Degree Centrality of a node $n$
$RD_{i,m}$	Relationship Degree of a node $i$ with respect to a specific buffered data message $m$
CGrAnt Global Metrics	
$ p $	Number of hops in a complete path $p$
$BU_{n,d}$	Betweenness Utility of a node $n$ in relation to $d$
CGrAnt Composite Metrics	
$BU_{J,d}^i$	Betweenness Utility of the social network of $i$ in relation to a destination $d$
$SP_{n,d}$	Social Proximity between nodes $n$ and $d$
$U_{n,d}$	Utility of a node $n$ as a message forwarder to $d$
$\eta_{n,d}$	Heuristic Function measured by $SP_{n,d}$
$\tau_{(i,y),d}$	Pheromone concentration on each link $(i,y)$ belonging to a path to node $d$
$SD_n$	Stagnation Degree of a node $n$
$SD_J^i$	Stagnation Degree of the social network of node $i$
$Q_{p_{s_m,d_m}^k}$	Quality of a path $p$ (from $s_m$ to $d_m$ ) constructed by FA $k$
CGrAnt Knowledge	
$Dom_i$	Domain Knowledge of node $i$
$His_{i,d}$	History Knowledge of $i$ with respect to $d$
$Sit_{i,m}$	Situational Knowledge of $i$ with respect to $m$

---

**Algorithm 1** Pseudo-code of the CGrAnt Message Forwarding.

---

```
1: Algorithm Initialization
2: for each message  $m$  in the buffer of node  $i$  do
3:    $best\_fwd_m \leftarrow \emptyset$ ; {No forwarder is assigned to  $m$ }
4: end for
5: { $BU$ , Pheromone concentration, and History knowledge are updated during the backward phase}
6: for each message  $m$  in the buffer of node  $i$  do
7:    $U_{i,d} \leftarrow U_{n,d}$ ; {Updating node  $i$  utility, as in Eq. 2}
8:   if ( $best\_fwd_m = \emptyset$ ) then
9:      $best\_fwd_{m,d} \leftarrow i$  {Initializing the Situational Knowledge};
10:     $U_{best\_fwd_{m,d}} \leftarrow U_{i,d}$ ; {Utility of the best forwarder for  $m$ }
11:   end if
12:    $search\_status\_exploration \leftarrow \text{true}$ ;
13:    $search\_status\_exploitation \leftarrow \text{true}$ ;
14:   {Domain Knowledge Influence}
15:   if ( $(RD_{i,m} \neq s_m) \text{ AND } (dr_i = 0)$ ) { $Dom_i = \text{medium } SD_i$ } then
16:      $search\_status\_exploration \leftarrow \text{false}$ ;
17:   end if
18:   if ( $(RD_{i,m} \neq s_m) \text{ AND } (dr_i = -1)$ ) { $Dom_i = \text{low } SD_i$ } then
19:      $search\_status\_exploration \leftarrow \text{false}$ ;
20:      $search\_status\_exploitation \leftarrow \text{false}$ ;
21:   end if
22:   

---


23:   Message Forwarding
24:   for all connections  $j$  do
25:      $U_{j,d} \leftarrow U_{n,d}$  {Updating node  $j$  utility, as in Eq. 2}
26:     if ( $(search\_status\_exploration) \text{ AND } (best\_fwd_m = i) \text{ AND } (U_{i,d} = U_{j,d})$ ) then
27:        $initial\_exploration \leftarrow \text{true}$ ;
28:     else
29:        $initial\_exploration \leftarrow \text{false}$ ;
30:     end if
31:     {Influence of History and Situational Knowledge}
32:     if ( $(search\_status\_exploitation) \text{ AND } (History\_Influence\_Function())$ ) then
33:       Forward  $m$  to  $j$  {History Knowledge Influence}
34:     else
35:       if ( $(initial\_exploration) \text{ OR } (Situational\_Acceptance\_Function())$ ) then
36:         {Situational Knowledge Update}
37:          $best\_fwd_m \leftarrow j$ ;
38:          $U_{best\_fwd_{m,d}} \leftarrow U_{j,d}$ ;
39:       end if
40:     end if
41:   end for
42: end for
43: if ( $best\_fwd_m \neq i$ ) then
44:   Forward  $m$  to  $j$  {Situational Knowledge Influence}
45: end if
46:  $History\_Influence\_Function()$ 
47: if ( $BU_{j,d} > His_{i,d}$ ) then
48:   Return TRUE { $dr_j = +1$ }
49: else
50:   Return FALSE { $dr_j = -1$  or  $dr_j = 0$ }
51: end if
52:  $Situational\_Acceptance\_Function()$ 
53: if ( $U_{j,d} > Sit_{i,m}$ ) then
54:   Return TRUE{Accept the solution  $j$ }
55: else
56:   Return FALSE
57: end if
```

---

Table 2: Simulation parameters.

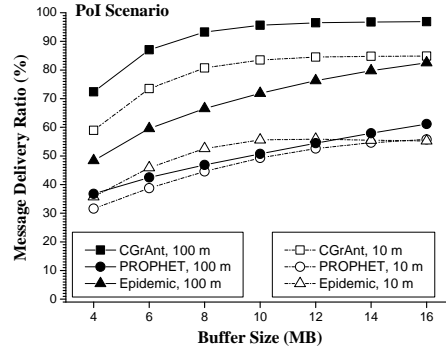
Protocol	Setting Parameters	Both scenarios	
CGrAnt	Pheromone evaporation rate	0.1	
	Heuristic function	$\{\mathbf{SP}_{n,d}, DC_n, BU_{n,d}, FB_n\}$	
	Pheromone concentration	$\{ p , DC_n, BU_n, \mathbf{DC}_n +  p \}$	
	Utility of a node	$\{\text{Heuristic}, \text{Pheromone}, \text{Heuristic} + \text{Pheromone}, \mathbf{Heuristic} + \mathbf{Pheromone} + \mathbf{RD}_{i,m}\}$	
PROPHET	Hop-count field (hops)	11	
	$P_{nic}$	0.75	
	$\gamma$	0.98	
	One time unit $Unit$ (s)	30	
	$\varphi$	0.25	
Epidemic	Hop-count field (hops)	11	
Protocol	Communication Parameters	PoI	WD
All	Number of nodes (N)	140	339
	Area ( $m^2$ )	8,800 x 7,800	10,000 x 8,000
	Nodes speed (m/s)	[0.5, 1.5]	[0.8, 1.4] (pedestrian), [7.0, 10.0] (car and bus)
	Waiting time (s)	100-200 (W1-Z1), 4000-5000 (W2-Z2)	300-500 (H), 10-30 (bus)
	Traffic generation rate (s)	50-90	100-150
	Message TTL (min)	{300, 600, 900, 1200, 1500, 1800, 2100}	{300, 600, 900, 1200, 1500, 1800, 2100}
	Nodes buffer (MB)	{4, 6, 8, 10, 12, 14, 16}	{4, 6, 8, 10, 12, 14, 16}
	Simulation time (s)	800,000	
	Warm up period (s)	5,000	
	Communication range (m)	10 (Bluetooth Devices), 100 (WiFi Devices)	
	Transmission rate (Mbps)	2 (Bluetooth Devices), 10 (WiFi Devices)	
	Number of simulations	30	

Table 3: Additive Analysis of the CGrAnt's Components

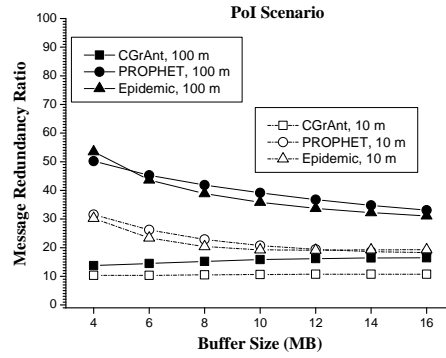
Configuration	(PoI    WD) Message Delivery %	(PoI    WD) Message Redundancy
1. Heuristic Function	46.26 $\pm$ 0.18    53.38 $\pm$ 0.60	18.36 $\pm$ 0.08    85.87 $\pm$ 0.94
2. Pheromone Concentration + RD	48.61 $\pm$ 0.20    54.97 $\pm$ 0.61	15.97 $\pm$ 0.08    68.62 $\pm$ 0.96
3. Domain Knowledge 1	49.40 $\pm$ 0.21    64.26 $\pm$ 0.63	15.11 $\pm$ 0.09    43.18 $\pm$ 0.41
4. Situational Knowledge	56.70 $\pm$ 0.27    61.04 $\pm$ 0.68	6.86 $\pm$ 0.04    10.87 $\pm$ 0.11
5. History Knowledge	58.93 $\pm$ 0.19    63.70 $\pm$ 0.69	10.37 $\pm$ 0.05    19.04 $\pm$ 0.16
6. Domain Knowledge 2	58.93 $\pm$ 0.19    63.19 $\pm$ 0.72	10.37 $\pm$ 0.05    12.43 $\pm$ 0.12

Table 4: Eliminatory Analysis of the CGrAnt's Components

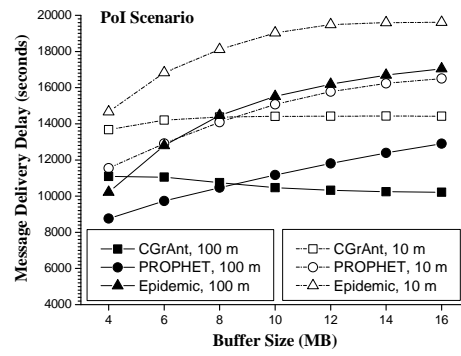
CGrAnt	(PoI    WD) Message Delivery %	(PoI    WD) Message Redundancy
All components	58.93 $\pm$ 0.19    63.19 $\pm$ 0.72	10.37 $\pm$ 0.05    12.43 $\pm$ 0.12
Without Domain Knowledge	57.90 $\pm$ 0.22    62.30 $\pm$ 0.67	10.45 $\pm$ 0.04    23.90 $\pm$ 0.28
Without Situational Knowledge	53.54 $\pm$ 0.22    64.23 $\pm$ 0.64	15.74 $\pm$ 0.07    43.32 $\pm$ 0.43
Without History Knowledge	56.70 $\pm$ 0.27    61.04 $\pm$ 0.68	6.86 $\pm$ 0.04    10.87 $\pm$ 0.11



(a) Delivery Ratio



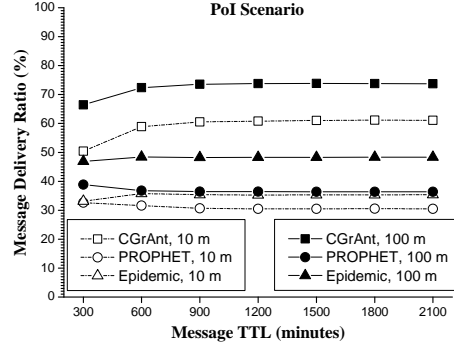
(b) Redundancy Ratio



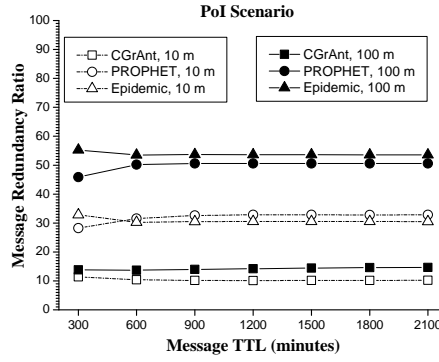
(c) Delivery Delay

Figure 6: Protocols' performance over different buffer sizes - PoI scenario.

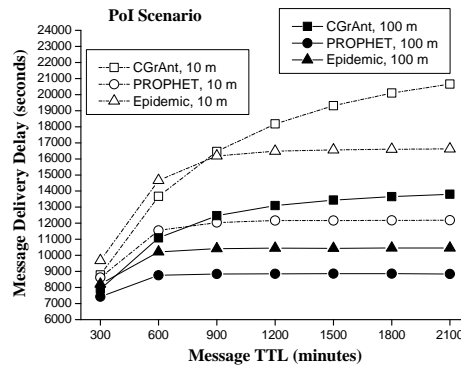




(a) Delivery Ratio

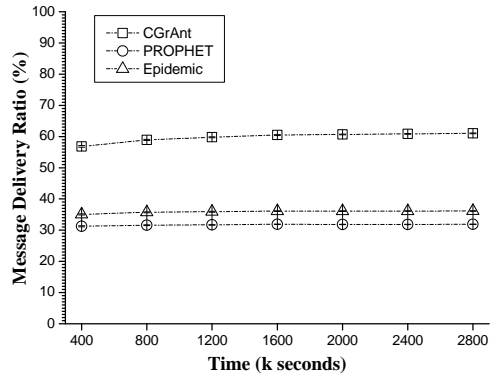


(b) Redundancy Ratio

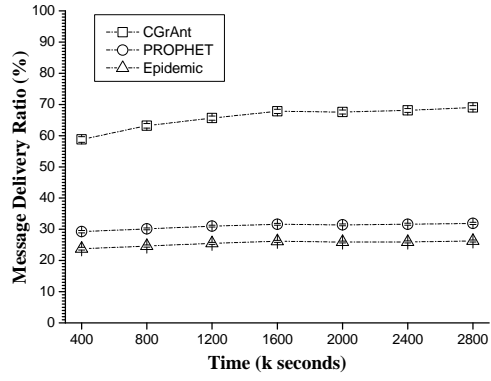


(c) Delivery Delay

Figure 7: Protocols' performance over different message TTLs - PoI scenario.



(a) Delivery Ratio - PoI scenario



(b) Delivery Ratio - WD scenario

Figure 8: Message Delivery Ratio over Different Simulation Time.

Table 5: Storage cost over different simulation time

Protocols/ Simulation Time	Registers/Bytes (100k sec.)	Registers/Bytes (200k sec.)	Registers/Bytes (300k sec.)	Registers/Bytes (400k sec.)	Registers/Bytes (500k sec.)	Registers/Bytes (600k sec.)	Registers/Bytes (700k sec.)
<b>CGrAnt</b>							
$FE_{i,d}$	93.97/1,127.65	116.01/1,392.17	126.73/1,520.74	131.67/1,580.06	134.79/1,617.43	136.51/1,638.17	137.54/1,650.51
$DE_{i,d}$	93.97/2,255.31	116.01/2,784	126.73/3,041.49	131.67/3,160.11	134.79/3,234.86	136.51/3,276.34	137.54/3,301.03
$DC_i$	2/16	2/16	2/16	2/16	2/16	2/16	2/16
$PT_i$ and $MS_i$	2/16	2/16	2/16	2/16	2/16	2/16	2/16
$SD_j$	90.82/1,453.14	116.01/1,856.23	126.73/2,027.66	131.67/2,106.74	134.79/2,156.57	136.51/2,184.23	137.54/2,200.69
$U_{best\_fwd_m}$	22.55/360.8	43.64/698.29	62.99/1,007.89	81.72/1,307.54	98.99/1,583.77	115.63/1,850.06	131.94/2,111.09
Pheromone Table	9.56/229.54	17.76/426.34	23.94/574.63	28.76/690.17	32.56/781.54	35.96/862.97	38.74/929.66
$BU_{i,d}$	8.58/102.94	16.1/193.2	22.04/264.43	26.76/321.09	30.54/366.51	33.88/406.54	36.74/440.91
$BU_{j,d}^i$	9.56/153.03	17.76/284.23	23.94/383.09	28.76/460.11	32.56/521.03	35.96/575.31	38.74/619.77
<b>PROPHET</b>							
Delivery Predictability List	139/2,224	139/2,224	139/2,224	139/2,224	139/2,224	139/2,224	139/2,224